

DOCKET FILE COPY ORIGINAL

EX PARTE OR LATE FILED



Suite 1000  
1120 20th Street, N.W.  
Washington, DC 20036  
202 457-3810

December 23, 1997

RECEIVED

DEC 23 1997

FEDERAL COMMUNICATIONS COMMISSION  
OFFICE OF THE SECRETARY

Ms. Magalie Roman Salas  
Secretary  
Federal Communications Commission  
1919 M. St., NW, Room 222  
Washington, D.C. 20554

RE: Ex Parte Presentation – Proxy Cost Models  
CC Docket No. 96-45

Dear Ms. Salas:

The attached materials on geocoding and customer location in the Hatfield Model, v5.0 are being provided to Brian Clopton of the Universal Service Branch of the Common Carrier Bureau in regards to his examination of cost models for universal service in CC Docket Nos. 96-45 and 97-160.

Two copies of this Notice are being submitted to the Secretary of the FCC in accordance with Section 1.1206(a)(2) of the Commission's rules.

Sincerely,

A handwritten signature in cursive script, appearing to read "Richard N. Clarke".

Richard N. Clarke

Attachments

cc: Sheryl Todd  
Brian Clopton

No. of Copies rec'd CH 4  
List ABCDE

## Example of Customer Location: Raw Address Files to Clustered Outputs

RECEIVED

DEC 23 1997

FEDERAL COMMUNICATIONS COMMISSION  
OFFICE OF THE SECRETARY

This document describes by example the process used by PNR and Associates to locate and cluster telephone customers for use by the Hatfield Model, v5.0.. It exhibits PNR's customer location process by following 200 addresses (100 residential and 100 business) heuristically through each step of the process – from raw data files to clustered outputs.

### Input Data

Customer location starts with raw data provided by two outside vendors. Metromail, Inc. provides residential information, and Dun & Bradstreet, Inc. provides business data. Both data sources provide PNR with their full address database for use in customer location.

For exposition, 200 actual addresses in the Gunnison, Colorado wire center are identified. The following tables show examples of both the residence and business datasets.

*Table 1 - Example of Residential Data*

145 TEOCALLI RD	CRESTED BUTTE	CO	81224
146 CACTUS HILL DR	GUNNISON	CO	81230
15 GLORIA PL	CRESTED BUTTE	CO	81224
151 UTE LN	GUNNISON	CO	81230
156 ZELIGMAN ST	CRESTED BUTTE	CO	81224
160 SUMMER HOME DR	CRESTED BUTTE	CO	81224
161 TOMICHI LN	GUNNISON	CO	81230
1611 STATE HIGHWAY 114	GUNNISON	CO	81230
166 TEOCALLI RD	CRESTED BUTTE	CO	81224
167 KINNIKINNICK LN	CRESTED BUTTE	CO	81224
17300N N STATE HIGHWAY 1	ALMONT	CO	81210
1747 STATE HIGHWAY 114	GUNNISON	CO	81230
186 TOMICHI LN	GUNNISON	CO	81230
18806 COUNTY ROAD 765	ALMONT	CO	81210
19015 COUNTY ROAD 765	ALMONT	CO	81210
200 INDUSTRIAL PARK RD	GUNNISON	CO	81230
213 RAINBOW RD	ALMONT	CO	81210
21946 STATE HIGHWAY 135	CRESTED BUTTE	CO	81224
23990 STATE HIGHWAY 135	CRESTED BUTTE	CO	81224

Table 1 shows an example of the residential address information used in customer location. Each record includes a street address, an apartment number (where indicated), city, state, and 5-digit zip code.

**Table 2 - Example of Business Data**

115 W GEORGIA AVE	GUNNISON	CO	812302421
117 TAYLOR	GUNNISON	CO	81230
118 COUNTY ROAD 13	GUNNISON	CO	812309702
12 SNOWMASS RD	CRESTED BUTTE	CO	81224
120 S IOWA ST	GUNNISON	CO	812302308
123 W TOMICHI AVE	GUNNISON	CO	812302345
124 N MAIN ST	GUNNISON	CO	812302331
127 ELK AVE	CRESTED BUTTE	CO	81224
129 ELK AVE	CRESTED BUTTE	CO	81224
130 COUNTY ROAD 742	ALMONT	CO	812109728
138 N MAIN ST	GUNNISON	CO	812302331
139 N MAIN ST	GUNNISON	CO	812302330
142 N MAIN ST	GUNNISON	CO	812302331
1445 STATE HIGHWAY 135	GUNNISON	CO	812309243
1516 COUNTY ROAD 10	GUNNISON	CO	812309601
1681 GOLD BASIN RD	GUNNISON	CO	81230
1730 COUNTY ROAD 744	ALMONT	CO	812109729
200 E VIRGINIA AVE	GUNNISON	CO	812302248

Table 2 shows an example of the business data used in customer location. Each record consists of a street address, any applicable suite or floor information, city, state, and 9-digit ZIP code (where this is not available, a 5-digit ZIP is included).

### **Geocoding**

The first step in customer location is to geocode each record in both datasets. Geocoding is used to accurately assign known customer addresses to actual, physical locations. It involves the assignment of latitude and longitude coordinates to actual street addresses. Geocoding also allows customer locations to be associated with their census block.

Initially, the address coding module in Centrus™ Desktop compares the customer's street address from the input file to the address records contained in the USPS ZIP+4 directory and GDT's enhanced street network files. If the address is matched in the USPS files, the address is standardized and a ZIP+4 is also returned. Continuing, if the address is also matched in the street network files, Centrus™ Desktop determines a latitude and longitude for the location. If the address is not found in the street network files, a latitude and longitude is provided for the ZIP code level. While ZIP-based location information is useful in many applications, it is not currently used in the Hatfield Model's customer location process.

Table 3 again shows a portion of the 100 residential records in this example. After geocoding, the record contains a number of new fields. These include a latitude and longitude, a census ID code (with varying levels of specificity), a match code and a location code.

**Table 3 - Example of Geocoding output for Residential data**

STREET	CITY	ZIP	LONG	LAT	CENSUS ID	MATCH CODE	LOC CODE
42 HUMMINGBIRD LN	ALMONT	81210	-106.486424	38.781180	08051963600110	S80	AS0
213 RAINBOW RD	ALMONT	81210	-106.488218	38.788311	08051963600110	S80	AS0
954 COUNTY ROAD 744	ALMONT	81210	-106.767130	38.749458	08051963800124	S80	AS1
2809 COUNTY ROAD 38	GUNNISON	81230	-106.940758	38.470850	08051963600288	S80	AS1
18806 COUNTY ROAD 765	ALMONT	81210	-106.480000	38.723500	0805196360001	S80	ZB9E
5866 COUNTY ROAD 62	GUNNISON	81230	-106.617400	38.439800	0805196360001	S80	ZB9E
299 ZUNI TRL	GUNNISON	81230	-106.936300	38.504200	0805196360002	S80	ZC7K
146 CACTUS HILL DR	GUNNISON	81230	-106.643400	38.467800	0805196360001	S88	ZC7K
200 INDUSTRIAL PARK RD	GUNNISON	81230	-106.559900	38.504100	0805196360001	S89	ZC7K
34 SLATE VIEW LN	CRESTED BUTTE	81224	-106.882500	38.807100	0805196380001	S88	ZT7J
17300N N STATE HIGHWAY 1	ALMONT	81210	-106.846000	38.664600	0805196360002	S82	ZT9F

For purposes of customer location, the latitude and longitude assignments are critical as is the census block ID. Actual point geocodes are returned populated to the 6<sup>th</sup> decimal place with an accompanying CB designation. For addresses geocoded to a lesser level of accuracy, the latitude and longitude are returned populated to the 4<sup>th</sup> decimal place and at best, a census ID specific to the CBG. These three fields, longitude, latitude, and census ID, will be the key inputs to the clustering process.

Although a required geocoding output, Match Codes are not crucial for customer location. Match Codes indicate how well an address matches to USPS data. Of greater importance is the last column in this example - the Location Code. Location Codes indicate the accuracy of the geocode. For the purposes of the Hatfield Model, v5.0, only those geocodes assigned at the 6<sup>th</sup> decimal location are used in customer location. Due to this purposeful dismissal, only a few of the possible location codes are relevant.

**Table 4 - Geocoding Location Codes**

<b>AS0</b>	Indicates a house range address geocode. This is the most accurate geocode available. It is the best location and is offset 50' to the correct side of the road.
<b>AS1</b>	Indicates a house range address geocode but the side of the street is unknown. In these instances, the Census Block to the left side of the point but no offset is assigned. The point is located directly on the street.
<b>AS2</b>	Indicates that the address was interpolated onto a TIGER segment that did not originally contain address ranges. The location is offset 50' to the correct side of the road.
<b>AS3</b>	Indicates that the address was interpolated onto a TIGER segment that did not originally contain address ranges. In these instances, the Census Block to the left side of the point but no offset is assigned. The point is located directly on the street.
<b>ZXnX</b>	Indicates a geocode based on ZIP code data. This is NOT a point geocode and therefore is not used in Customer Location.

Any data points that do not have a location code beginning with the letter A are not used in customer location. These points, along with their accompanying data are excluded from further use.

The 200 example points are geocoded. In general, geocoding to the actual point location (i.e., 6<sup>th</sup> decimal place) is successful 70 to 80% of the time. Failure to geocode to this precision is due primarily to incomplete or incorrect road addresses, incorrect or incomplete underlying road or TIGER data, and rural addressing. Certain software enhancements will improve the accuracy of incomplete addresses while tightened postal requirements and E911 expansions will decrease the number of actively used rural route addresses.

Of the 100 residential points from the rural Gunnison wire center, 77 records were matched to the AS0 level, the best match possible. An additional 9 records were located to the proper point on the street, although no 50' offset was identified. The remaining 14 locations were geocoded to their 5-digit ZIP code. This leaves the customer location process with 86 usable residential points. For this set of sample data, business addresses do not geocode quite as successfully. Of the 100 business points, 60 records achieved an AS0 match. Eleven additional records were located to the proper point on the street with no offset from the road. This leaves the customer location process with 71 useable business point locations. Of the remaining 29 business records, 11 were geocoded to the 5 digit ZIP code level, while 18 records could not be matched at all. As Table 5 shows, where the address is only located to the ZIP code, Centrus applies a 4 decimal place lat/long combination, and when the address cannot be matched at any level, the latitude, longitude, and census ID cell remain empty.

**Table 5 - Examples of Geocoding output for Business Data**

STREET	CITY	STATE	ZIP	PLUS4	LONG	LAT	CENSUS_ID	MATCH CODE	LOC CODE
710 S 9TH ST	GUNNISON	CO	812303910	3910	-106.938124	38.537994	08051963700541	S90	AS0
720 N 8TH ST	GUNNISON	CO	812303518	3518	-106.939493	38.55109	08051963700305A	S90	AS0
118 COUNTY ROAD 13	GUNNISON	CO	812309702	9702	-106.930754	38.573011	08051963600496	S90	AS1
127 ELK AVE	CRESTED BUTTE	CO	81224		-106.987328	38.8696	08051963800318	T00	AS1
114 BUTTE AVE	CRESTED BUTTE	CO	81224		-106.987803	38.873244	08051963800311	T09	AS3
663 COUNTY ROAD 13 3	GUNNISON	CO	81230	8722				E024	E
6748 COUNTY 742	GUNNISON	CO	81230					E020	E
280 CEMENT CREEK RD	CRESTED BUTTE	CO	812249600	9600	-106.8837	38.8095	080519638001	S90	ZB9E
1730 COUNTY ROAD 744	ALMONT	CO	812109729	9729	-106.7681	38.7421	080519638001	S90	ZC9G

While the geocoding software offers many more additional elements for output, only the following addressing elements are used in Customer Location:

Address  
City  
State  
ZIP  
ZIP+4  
Latitude  
Longitude  
Census Block  
Match Code  
Location Code

### **Determining the target number of Residence Locations**

The number of geocoded points is used to determine the number of residential locations. Since the residential database is based on a mailing list, duplicate records must be eliminated. Records are considered to be duplicates if they have identical lat/long assignments, as well as the same phone number. Where this repetition occurs, one of the records is deleted. Only those records that have been geocoded to the 6<sup>th</sup> decimal place are used. As explained in the the Hatfield Model 5.0 Description document, in those instances where Metromail counts exceed Claritas in a CBG, the Metromail data is used as the total residential count. In those CBGs where Claritas exceeds Metromail, Claritas is used as the CBG total. In these instances, the Claritas excess is distributed proportionally to the CB household counts based on 1990 Census distributions.

### **Determining the target number of Business Locations**

While the Dun & Bradstreet database accounts for approximately 11 million businesses nationally, it is believed to miss an additional 1 million businesses. Business surrogate points are created for the D&B businesses that could not be geocoded, and for the additional 1 million businesses not in D&B's database. Unlike residential data, where there are Census counts and other private projections, there is no other source to which business data can be adjusted. Given this, unlocated businesses are pooled together and distributed nationally to CBs based on identified business locations.

### **Creating Surrogate or Gross-Up Points**

Now that the target number of customer locations by CB have been determined, any excess, or "surrogate," points need to be physically located to their respective census blocks so that they may be clustered. "Unlocated" customers are assumed to be located uniformly along the periphery of the CB, and the geocodes implied by these placements are added to the customer location file. This addition of surrogate points assures that the customer location file contains geocodes for 100% of the target number of customer locations by CB.

**Figure 1 - Point Locations in Gunnison CO**

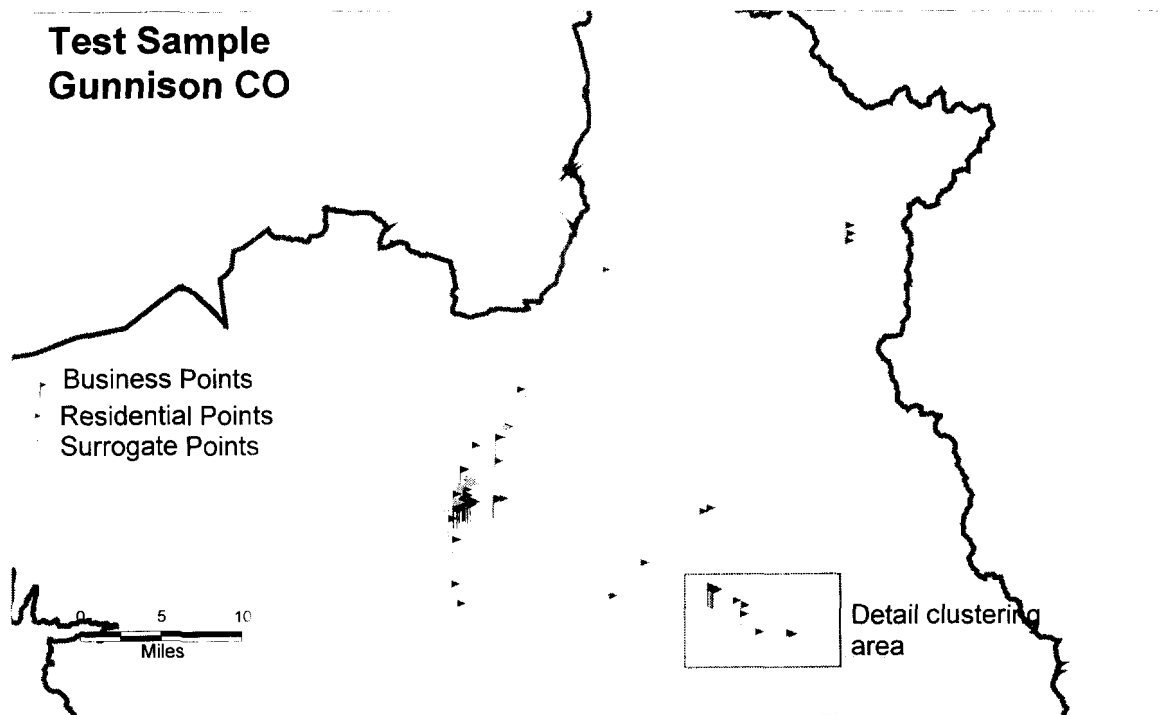


Figure 1 shows the portion of the Gunnison, Colorado wire center (GNSNCOMA) used in this example. The pink outline indicates the wirecenter boundary while the blue outlines indicate CBG boundaries. As identified in the legend, the blue flags show the locations of the 71 business locations and the red flags show the locations of the 86 residential locations. The gray flags are surrogates which represent the total number of surrogates, both business and residential. Once all points have been located, clustering may begin.

### **Clustering**

In order to determine efficient engineering of telephone plant, each WC's collection of customer locations is analyzed to determine what cluster patterns exist. While many off-the-shelf routines were trialed, no existing algorithm supported the unique needs of telephone customer location. The clustering algorithm developed by PNR ensures that identified cluster configurations comport with telephone plant engineering specifications.

**Figure 2 - Customer Location Clustering Methodology**

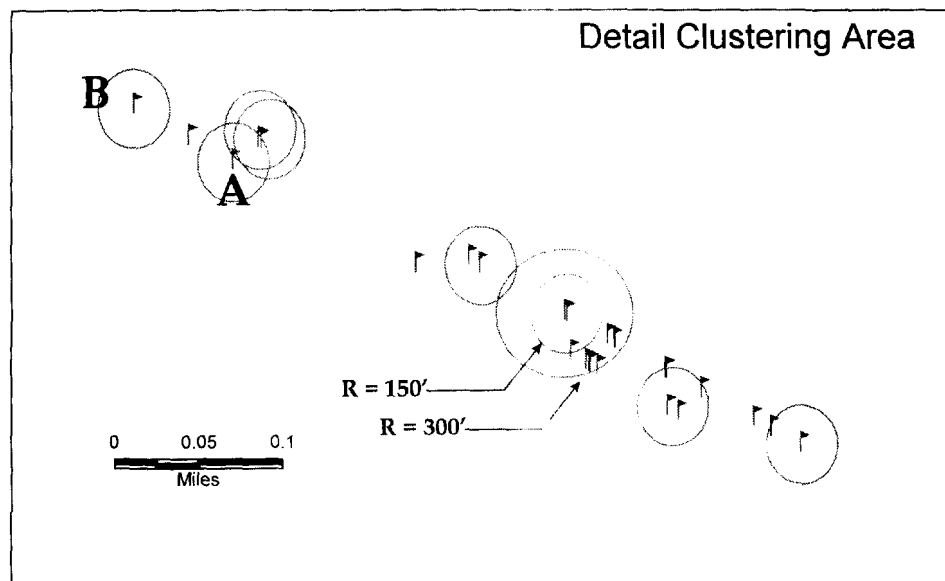


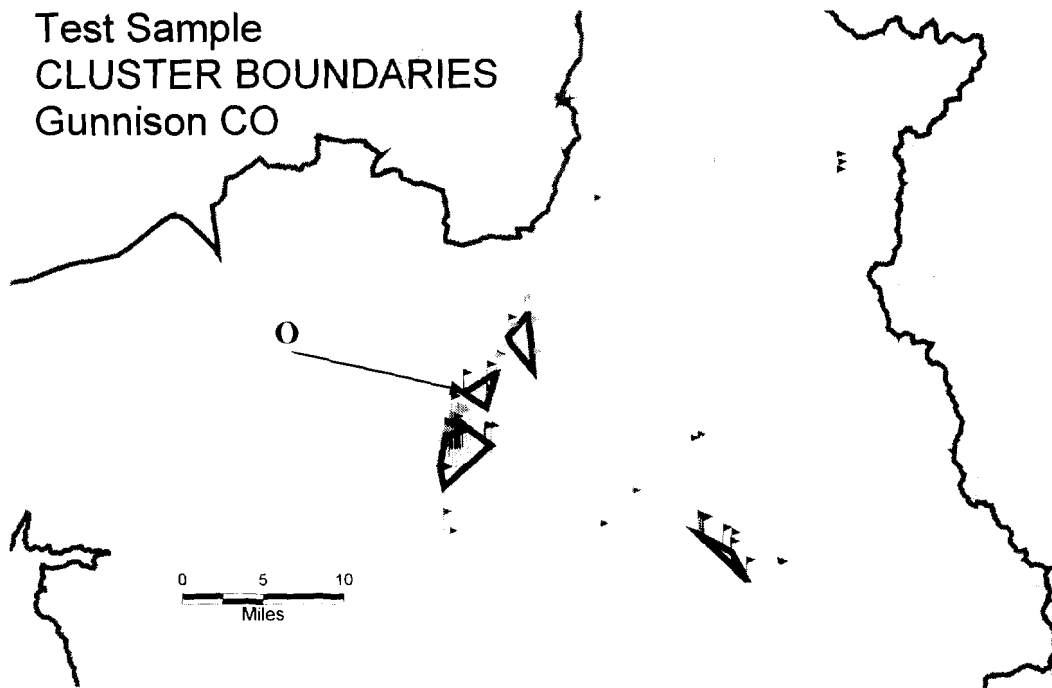
Figure 2 shows a subset of the Gunnison example (this subset is highlighted in Figure 1). Clustering is an iterative process which can begin at any Point A. A circle with a radius of 150' is drawn around Point A. Any points that falls within the circle are automatically considered to be part of the cluster. Continue with other points in the cluster, continually drawing 150' circles until no new points fall within any of the radii. Move to any Point B which is not part of an existing cluster and repeat the process until all points have been considered. Return to Point A and create a circle 300' in radius. The process continues in a iterative manner until all points have been clustered or until one of the engineering criteria have been met.

ENGINEERING REQUIREMENTS FOR CLUSTERS	
No point in a cluster may be more than 18 kft. distant from the cluster's centroid	
No cluster may exceed 1800 lines in size	
No point in a cluster may be farther than 2 miles from its nearest neighbor in the cluster	
A main cluster must contain at least 5 points, smaller clusters are called outlier clusters	

Figure 3 shows the example data. The series of red polygons around groups of flags are the cluster boundaries. All the points within a cluster meet all the engineering criteria as defined by the user.

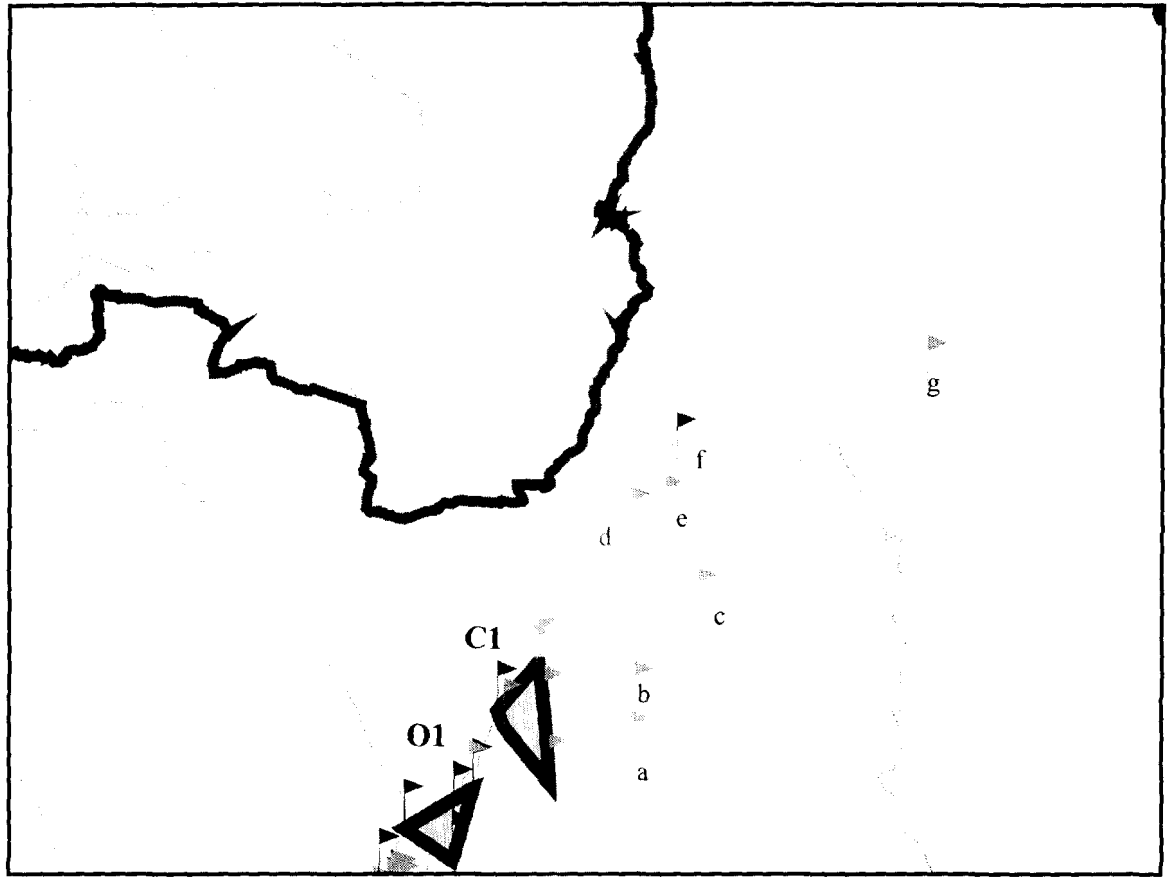


*Figure 3 - Clusters formed from Gunnison CO example data*

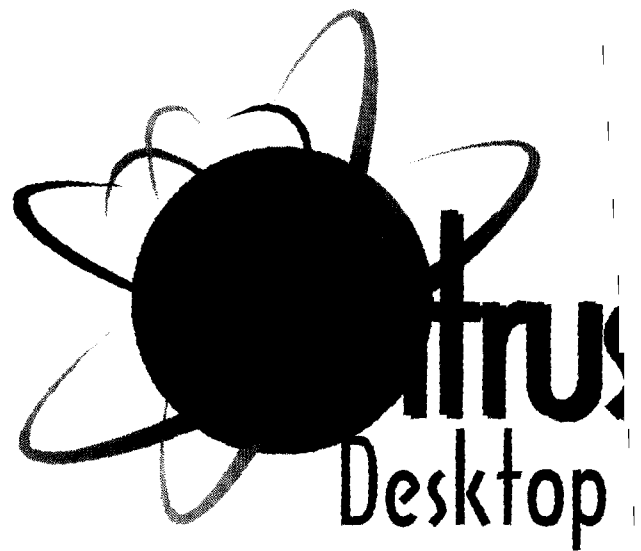


In Figure 3 is an arrow pointing to cluster O. Cluster O represents an “Outlier” cluster. These outlier clusters are then attached to a main cluster via a “chaining” process. The rectilinear distance between all clusters, whether between two outlier clusters (an outlier cluster may encompass anywhere from 1 to 4 points) or between an outlier and main cluster is calculated. The minimum distance in this full matrix of point is found and linking begins there. Each outlier cluster is linked to a main cluster either directly or through a chain of other outliers.

**Figure 4 - Detail of "outlier" clusters**



The shortest distance evident in Figure 4 is found between *outlier clusters d and e*. These two *clusters* are linked. The next shortest distance is between *outlier clusters a and b* -- these two *clusters* are linked. Continuing, the next shortest distance is between *e and f*, thus *f* is linked to *d and e*. *A* is now closest to main cluster **C1** and thus, *a and b* are linked to a main cluster. *Cluster c* is linked to *clusters e, d, and f* and is then linked to *a and b*. Now, *clusters c, d, e, and f* are linked to main cluster **C1** via *a and b*. Finally, *cluster g* is linked to the rest of the chain.



## **User's Guide**

*Version 2.01*



**Qualitative Marketing Software, Inc.**

# **Centrus™ Desktop User's Guide**

*Manual Version: 1297*

Copyright © 1994, 1995, 1996, 1997 Qualitative Marketing Software, Inc. All rights reserved.

National ZIP+4 Directory © United States Postal Service 1993, 1994, 1995, 1996, 1997

## **Trademark Notices:**

ZIP Code and ZIP+4, are registered trademarks of the U.S. Postal Service. *FASTforward* is a service mark of the U.S. Postal Service.

Within this manual, other trademarked names may be used. Rather than putting a trademark symbol in every occurrence of a trademarked name, we state that we are using the names only in an editorial fashion, and to the benefit of the trademark owner, with no intent of infringement on the Trademark.

## **USPS Notice:**

Qualitative Marketing Software, Inc. holds a non-exclusive license to publish and sell ZIP+4 databases on optical and magnetic media. The price of the QMS product is neither established, controlled, nor approved by the U.S. Postal Service.

# Contents

## Chapter 1

<b>About Centrus Desktop .....</b>	<b>1</b>
What Centrus Desktop Can Do For You .....	1
An Overview Of Centrus Desktop Components.....	2
Data Locator Module .....	2
Address Coding Module .....	2
CentrusCOA Module.....	2
Demographic Coding Module.....	3
Spatial Coding Modules.....	3
Putting It All Together.....	3
Minimum System Requirements.....	4
About This User Guide .....	4
Contacting Technical Support .....	5

## Chapter 2

<b>An Introduction to Centrus Desktop.....</b>	<b>7</b>
Getting Started .....	7
Start Centrus .....	7
Import Point-in-Polygon and Closest Site Layers.....	8
Setting up a Task.....	10
Open Input and Output Tables.....	10
Specify Data Locator Inputs and Outputs.....	11
Specify Address Coding Inputs and Outputs .....	12
Specify Demographic Inputs and Outputs .....	13
Specify Point-in-Polygon Inputs and Outputs.....	14
Specify Closest Site Inputs and Outputs.....	15
Processing a Task.....	16

## Chapter 3

<b>Configuring Centrus Desktop.....</b>	<b>19</b>
Installing Centrus Desktop.....	19
Improving Performance .....	19
License Files .....	20
Configuring Centrus Desktop .....	20
Specifying File Paths .....	21
Setting Up Tasks.....	22
Creating a New Task.....	22

Saving a Task .....	22
Opening a Task.....	22
Editing a Task.....	22
Renaming a Task.....	23
Setting Options .....	23
Address Coding Options Tab .....	23
Data Locator Options Tab .....	28
Point-in-Polygon Options Tab.....	31
Closest Site Options Tab .....	32
Verifying Your Settings .....	34
Verify Current .....	34
Command Line Syntax .....	34

## Chapter 4

<b>Centrus Desktop Tools .....</b>	<b>37</b>
QuickFind .....	37
Data Locator Mode .....	37
Standard Mode .....	38
Multiline Mode .....	38
Address Wizard Mode .....	39
The Query Function .....	41
Data Viewer.....	47
Navigating in Data Viewer.....	47
Map Viewer.....	48
Manual Geocoding With the Map Viewer .....	50
Working With Layers .....	50

## Chapter 5

<b>Specifying Files in Centrus Desktop .....</b>	<b>53</b>
File Types Used in Centrus Desktop.....	53
The Tables Tab .....	54
Specifying An Input Table .....	54
Specifying An Output Table.....	55
Specifying A Reject File .....	55
Report Files .....	56

## Chapter 6

<b>The Data Locator Module.....</b>	<b>57</b>
About the Data Locator Module.....	57
How Data Locator Works.....	57
Specifying Data Locator Input Fields.....	58
Process if Field is Blank .....	58
Assigning Data Locator Output Fields .....	58
Assigning Field Names.....	59
Tips on Using Data Locator .....	59
Decide Which Information You Want to Extract.....	59
Define a Task to Extract the Information .....	60
Evaluate Results and Redefine the Task .....	60

Available Data Locator Elements .....	61
---------------------------------------	----

## Chapter 7

<b>The Address Coding Module .....</b>	<b>65</b>
About the Address Coding Module .....	65
Area Coverage .....	65
How Address Processing Works .....	65
Using the Address Coding Module .....	66
Specifying Address Input Fields .....	66
Using Different Input Formats .....	67
Multiline Input .....	67
General Guidelines for Using Multiline Mode .....	68
Process if Unmatched or Field is Blank .....	68
Optimization Sort .....	68
Assigning Address Output Fields .....	68
Assigning Address Elements .....	69
Creating New Fields .....	69
Unassign .....	69
Standardization Defaults .....	69
Geocode Defaults .....	70
Blank if Unmatched .....	70
Available Address Data Elements .....	71

## Chapter 8

<b>The CentrusCOA Coding Module .....</b>	<b>75</b>
About the CentrusCOA Module .....	75
Configuring the CentrusCOA Module .....	75
Specifying CentrusCOA Input Fields .....	76
Optimization Sort .....	77
Assigning CentrusCOA Output Fields .....	77
Assigning Field Names .....	77
Standardization Defaults .....	78
Geocode Defaults .....	79
Available CentrusCOA Elements .....	79

## Chapter 9

<b>The Demographic Coding Module .....</b>	<b>85</b>
About the Demographics Module .....	85
Specifying Demographic Input Fields .....	85
Process if Field is Blank .....	86
Assigning Demographic Output Fields .....	86
Assigning Field Names .....	86
Available Demographics Elements .....	87

## Chapter 10

<b>The Point-in-Polygon and Closest Site Modules .....</b>	<b>89</b>
About the Spatial Analysis Modules .....	89
Object Files and Layers .....	89

Point-in-Polygon Analysis .....	90
Specifying Point-in-Polygon Input Fields .....	90
Assigning Point-in-Polygon Output Fields .....	90
Closest Site Analysis .....	92
Specifying Closest Site Input Fields .....	93
Process if Field is Blank .....	93
Assigning Closest Site Output Fields .....	93

## Chapter 11

<b>Processing .....</b>	<b>95</b>
Processing a Data File .....	95
Batch Processing .....	95
Interactive Processing .....	95
The Process Dialog .....	96
Possible Matches Dialog .....	98
Processing with CentrusCOA .....	99
Customer Information .....	100
Matching Logic .....	100
Update Records Matched To .....	100
Dates .....	100

## Reference

<b>System Messages and Codes .....</b>	<b>101</b>
Match Codes .....	101
Return Codes For Match Codes—First Hex Digit .....	102
Return Codes For Match Codes—Second and Third Hex Digit .....	102
Location Codes .....	103
Address Location Code Descriptions .....	103
ZIP+4 Centroid Location Code Descriptions .....	104
ZIP Centroid Matching .....	107
<b>Customizing the Data Locator Module .....</b>	<b>109</b>
Regular Expressions .....	109
More About Regular Expressions .....	110
Custom Dictionaries In Data Locator .....	111
Creating A Custom Dictionary File .....	111
Examples Of Custom Dictionary Entries .....	114
<b>Premium Demographics .....</b>	<b>115</b>
Claritas Demographics .....	115
National Decision Systems Demographics .....	119
<b>GSD Split for Windows .....</b>	<b>123</b>
About GSD Split For Windows .....	123
Installing GSD Split For Windows .....	123
Starting GSD Split For Windows .....	123
Using GSD Split For Windows .....	124
Default Path Settings .....	124
Select Extract Files Section .....	124
<b>Preparing ASCII Text Files .....</b>	<b>127</b>
About Format Files .....	127
Delimited File Example .....	128



---

Fixed Length File Example .....	129
<b>CASS Report and Bulk Rate Information.....</b>	<b>131</b>
Cass (3553) Report Information .....	131
Bulk Rate Requirements And Guidelines .....	132
Where To Get Help .....	132
<b>Audit and Log Report Samples.....</b>	<b>135</b>
Audit Report.....	135
Log Report .....	135
<b>FASTforward Reports.....</b>	<b>139</b>
About FASTforward Reports.....	139
Monthly Report .....	139
NDI Summary .....	140
Service Log .....	140
KeyFile.....	141
General .....	141
<b>Glossary.....</b>	<b>143</b>
<b>Index.....</b>	<b>147</b>



# *Chapter 1*

## **About Centrus Desktop**

### **What Centrus Desktop Can Do For You**

Your organization spends thousands of dollars each year collecting the names and addresses of your customers and prospects. Unfortunately, collecting this information is only part of the battle. In today's aggressive marketplace you must know more about your customers to do a better job of targeting them. Centrus Desktop helps you learn more about your customers by properly fielding and cleansing your data and adding valuable information, such as:

- Age range
- Income range
- Ethnicity
- Gender
- Presence of children
- Latitude and longitude
- Census codes
- Postal codes
- Distance to user-defined locations

Centrus Desktop correctly fields and parses, cleans, standardizes, geocodes, spatially codes, and demographically codes your data *in a single pass*. It can use either interactive or high speed batch processing, so there's no need to tie up your mainframe. Centrus Desktop works natively with a variety of data formats, including text (delimited and fixed-length ASCII), dBase, Access, Excel, FoxPro, and Paradox. Using ODBC, Centrus Desktop can work with Oracle, SQL Server, and other data formats, eliminating the need to translate files into a different format for use in Centrus Desktop.

With the CentrusCOA module and the licensed USPS *FASTforward* system, you can quickly update addresses, reduce undeliverable mail, and maximize your postage discounts. Centrus Desktop exceeds U.S. Postal Service CASS (Coding Accuracy Support System) requirements, making your next mailing eligible for substantial postal savings and guaranteeing accurate address standardization.

## An Overview Of Centrus Desktop Components

Centrus Desktop uses a flexible, modular design that lets you license any or all of the components your organization needs. You can choose among the Data Locator module, Address Coding module, Demographic Coding module, and Spatial Coding modules (performing both Point-in-Polygon analysis and Closest Site analysis).

### Data Locator Module

The Centrus Data Locator module scans your data in freeform or fielded text and identifies name elements, address elements, and a user-defined expression (such as e-mail address, phone number, customer number, or Social Security number). It then quickly and accurately outputs the information to fields of your choice.

The Data Locator also identifies ethnicity and gender, and can generate appropriate salutations for your mailings. It includes a user-extensible dictionary which allows you to add additional name information, tailoring the Data Locator to your specific needs.

### Address Coding Module

The Centrus Address Coding module cleanses addresses to the highest level, exceeding the U.S. Postal Service CASS standards. Since Centrus Desktop is CASS Certified, you're assured that mailings will be correctly delivered. With over 20 million addresses changing every year, our bimonthly updates ensure the integrity and effectiveness of your valuable address data.

As Centrus Desktop corrects addresses, it incorporates a comprehensive address-level geographic coding function specifically designed to spatially enable your data. This process adds key components to each database record including latitude, longitude, Census tract, block group, and more.

### CentrusCOA Module

Effective July 1, 1997, the rules governing discounts on automated First Class mail require you to update lists with current USPS change-of-address information prior to mailing. Until now, move update information was only available from a small group of NCOA providers licensed by the USPS. But thanks to the Postal Service's new *FASTforward* technology initiative, your company can have access to current change-of-address data without the turnaround delays and expenses associated with using an NCOA service bureau.

CentrusCOA, combined with the Centrus Address Coding Module, is an effective, easy-to-use way to ensure that your mailing lists are always current. Here's how it works:

First, the Address Coding module corrects addresses to meet USPS CASS standards. The CentrusCOA module then submits the addresses to the dedicated USPS *FASTforward* computer for move update processing, matching name and address information in your file against the current

National Change Of Address (NCOA) data. The *FASTforward* change-of-address files are always current, updated automatically by the USPS every week.

## Demographic Coding Module

The Centrus Demographic Coding module appends valuable demographic variables and cluster codes to your existing databases. Knowing your customers' habits and lifestyles is critical to understanding them. Once you identify who your customers are, you can use this information to target your marketing efforts to similar groups or households, increasing your chances of reaching the correct audience.

Standard 1990 Census Demographics are included as a basic feature of the Demographic Coding module. Qualitative Marketing offers current year and five-year projected block group demographics and cluster codes at a ZIP+4 level from consumer demographic firms including National Decision Systems and Claritas. Cluster codes provide a simple way to identify, understand, and target consumers. They are widely used for strategic planning, customer profiling, direct mail, site analysis, media planning, and many other functions. Centrus Desktop appends demographic codes based on each address's geographic location, using either Census block group or ZIP+4.

## Spatial Coding Modules

The Centrus Spatial Coding modules give you extremely fast, flexible processing for the most frequently used GIS functions: *Closest Site* analysis, and *Point-in-Polygon* analysis. Closest Site analysis is a highly efficient distance test between two sets of points. The healthcare, banking, and insurance industries use this type of analysis to match up doctors and patients, find the nearest bank branch to a customer, and calculate the distance between a home and a fire station.

Point-in-Polygon analysis lets you accurately determine which areas a given point falls within. Retail, insurance, and telecommunication industries use Point-in-Polygon to identify key markets, insurance risk areas, and telephone service areas.

## Putting It All Together

Centrus Desktop is the integrated solution for your data quality, data enhancement and data analysis needs. You simply specify a database, configure each module to deliver the information you want, and start processing. Centrus Desktop saves your preferences as a "task" and lets you choose whether to process interactively or in batch mode. You can run the same task on different databases, or configure different tasks for different purposes. You can also process single addresses interactively using Quick Find.

## Minimum System Requirements

Centrus Desktop has the following minimum system requirements:

- A computer running Windows 95 or Windows NT 4.0 (Intel).
- Pentium processor or higher.
- 16 megabytes of memory.
- A hard disk with a minimum of 100 megabytes of free disk space.
- 3 1/2" disk drive.
- CD-ROM drive.
- Mouse or other pointing device.

A fast disk drive, fast processor, and more memory will improve performance. Hard disk space requirements vary with the Centrus Desktop modules and data files licensed.

## About This User Guide

This manual explains the functionality and use of all Centrus Desktop components. "An Introduction to Centrus Desktop" on page 7, contains a tutorial which will walk you through the basics of using Centrus Desktop. "Configuring Centrus Desktop" on page 19, "Centrus Desktop Tools" on page 37, and "Processing" on page 95 contain information common to all components. To learn more about the Address Coding, Spatial Coding, Demographics, or Data Locator modules, locate the desired topic in the table of contents.

To help distinguish text that appears on menus or dialogs, such text is displayed in **bold** or *italics*. Commands on menus and buttons are shown in **bold**, as are functions and dialog box options. *Italics* are used when referring to a section of a dialog in a description.

The manual assumes that you are familiar with Windows and that you know how to use a mouse and choose dialog options. If you are not comfortable with Windows, you might want to take the introductory tour:

1. Click the **Start** button on the Taskbar. Windows will open the start menu.
2. Click on **Help**. The Windows help file will open.
3. If the Contents tab is not selected, click it now.
4. Double-click on either "Tour: Ten Minutes to Using Windows" (Windows 95) or "Introducing Windows NT" (Windows NT).

## **Contacting Technical Support**

If you have any technical questions regarding the use of Centrus Desktop, please contact Qualitative Marketing Software's development office:

Qualitative Marketing Software, Inc.  
2900 Center Green Court South  
Suite 200  
Boulder CO 80301-5418

(303) 442-2838 (voice)

(303) 440-3523 (fax)

[support@qmsoft.com](mailto:support@qmsoft.com)

<http://www.qmsoft.com>





# Chapter 2

## An Introduction to Centrus Desktop

### Getting Started

This chapter is a tutorial designed to show you how to use each of the modules in Centrus Desktop. First, you'll install Centrus and configure it to analyze some sample data files. Then, you'll set up a "task" specifying how you want each module to process the data. Finally, you'll process the task and view the results.

Your software package includes three CD-ROMs—one installation disc, and two data discs.

To install Centrus Desktop, simply insert the installation disc into your CD-ROM drive and wait several seconds. Windows should automatically start the QMSetup program. If the "Autorun" feature is disabled on your machine, start QMSetup by selecting **Run** from the Windows Start menu, then typing: **D:\QMSetup.exe**. (Replace the letter "D" with the letter of your CD-ROM drive.)

Follow the on-screen prompts to complete the installation. If you are updating from an earlier version, the setup program will recognize this and prompt you with the recommended action. You may also be prompted to insert the diskette containing your license file.

For additional information on installing Centrus Desktop on your platform, refer to "Configuring Centrus Desktop" on page 19 and the installation notes that came with the CD-ROMs.

### Start Centrus

From the Windows Start menu, select **Centrus 2.0 Demo**. This will allow you to use all of the modules except CentrusCOA to process up to 25 records—even if you are only licensed for Centrus Address Coding Module.

